Maastricht University

Time series modelling in repeatedly conducted sample surveys

Citation for published version (APA):

Balabay, O. (2016). *Time series modelling in repeatedly conducted sample surveys*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. https://doi.org/10.26481/dis.20160511ob

Document status and date:
Published: 01/01/2016

DOI:
10.26481/dis.20160511ob

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 05 May. 2023

Valorisation

The well-being of a country is defined to a large extend by good governing and planning, which, in turn, is not possible without reliable quantitative and qualitative data. Collecting, summarizing and releasing such data about the country's state of affairs is usually laid over the shoulders of the country's national statistical institute (in the Netherlands, it is Statistics Netherlands, or Centraal Bureau voor de Statistiek (CBS)). If to be described in just a couple of words, the matter of this thesis is improving the quality of official statistical figures. The EUROSTAT defines several quality dimensions (EUROSTAT (2014)), of which two are considered in this work: 1. accuracy and reliability and 2. coherence and comparability.

The first problem arises because observing the whole population, i.e. census, can be very costly or even impossible and is thus hardly ever done in statistical practice. Instead, a sample is drawn that should be representative of the whole population. It is obvious that the point-estimate for the quantity of interest (e.g., for the true average or total) based on this sample is not exactly equal to the true value in the population, and drawing another sample would result in a different estimate. This variability of point-estimates across all possible samples is summarized by the (design-based) variance or standard error. The bigger the sample size is, the closer we expect the estimate to be to the true value, and hence, the smaller the variance is. High design-based variances can be detected in time series with the naked eye, as they result in excessively volatile point-estimates.

Sometimes, estimates for small areas or domains are required (and the demand on these has been growing in recent years), where the sample size may be so small (could be even zero) that the estimated variance is unacceptably large. Think, for instance, of an estimated unemployment rate of $3(\pm 4)\%$. Such a huge uncertainty around the point-estimate tells us virtually nothing about the existing unemployment rate. But even if it is estimated more precisely up to ±1%, it is still not precise enough, since a two-percent uncertainty around the unemployment rate makes it difficult for policy makers to decide on their fiscal and monetary instruments. The confidence interval bands of ±0.1%, as one usually sees in unemployment statistics produced by EUROSTAT, are achievable but require interviewing much more people. This means an increase in public money expenditures due to additional interviewing, as well as in the total administrative burden respondents experience by participating in the survey. The problem with comparability stems from the fact that surveys sooner or later get redesigned due to cost of quality considerations. A change in a way data are collected renders the figures before and after a redesign incomparable over time. As one can see, NSIs produce estimates contaminated with different kinds of errors, while being a respectable source of data for policy makers or non-governmental researchers who usually treat these estimates as the true data in their econometric models or other kinds of analytics. Therefore, it is of utmost importance that data produced by NSIs are of best possible quality. Nowadays, when NSIs' budgets are shrinking and when the society is becoming ever more averse to administrative burden they experience from their NSIs, increasing the sample size is not the best of even feasible option, so statisticians have to look for more sophisticated methods to meet both constraints. In fact, NSIs dispose of lots of information that can be used in improving newly-produced figures without resorting to trivial sample size increases. It is obvious that information on the same quantities of interest from the past or from similar domains can be of a great use, but is usually discarded, though. With this dissertation, I have tried to raise the awareness about the potential of time series modelling (namely, state space and multilevel time series) techniques that address both problems of small sample sizes and incomparability. Committing to these techniques would put the production of official statistics on the rails of model-based, rather than design-based approach. Most NSIs in the world are, however, conservative in this respect and prefer relying on traditional (assumption-free) methods based on the survey design in production of their figures. This is understandable, because a bias-free implementation of new methods requires a careful model selection, while the penetration of ideas about the true power of these methods among official statisticians is not high. The knowledge transfer from academia to NSIs is limited because it often requires routine work from academics, as well as certain degree of expertise in survey methodology. In order to fill this gap, Statistics Netherlands tries to strengthen its collaboration with academia by establishing professorships and funding PhD research projects, like the present one. This dissertation is aimed at illustrating how information accumulated over time and space can be exploited in repeatedly conducted surveys to improve the quality of official statistical figures in terms of precision and comparability over time. As this research is conducted thanks to and in the interests of Statistics Netherlands, the objective has been to develop suitable time series models for several data sets for repeatedly conducted surveys with several survey redesigns. Variance reduction runs as a red thread through the whole dissertation, and in Chapters 2 and 5 it is the central issue. Chapter 2 applies the state space model-based approach to the Dutch Road Transportation survey that features several survey redesigns and insufficiently large sample sizes. It is shown that most variance reduction comes from borrowing information over time, i.e. even when domains are modelled individually. Jointly modelling all the domains in one model by exploiting the correlation between them, i.e. borrowing information over space, further reduces the variance of the newly produced time series. The resulting standard errors of the series are eventually 40 to 70 percent smaller than the design-based ones, with smaller domains experiencing more variance reduction compared to the larger ones. The coherence between estimated domain series, on the one hand, and the series that constitutes the sum thereof, on the other hand, is addressed in Chapter 3 based on the same survey. State space models are usually estimated with the help of the well-known to engineers Kalman filter. The problem with these models, despite their indisputable power, is that in practice the Kalman filter relies on estimated hyperparameters, rather than on their true values. Since there must always be a certain amount of uncertainty around any estimate, it should be taken into account; otherwise, the estimated variances of the newly produced series will be negatively biased (underestimated). The literature knows several methods to correct for the bias in variances estimated with the Kalman filter. However, before any newly developed theory can be applied in the regular production process, its value added and flaws must be verified. Usually, the performance of new methods is illustrated with simple examples where the superiority of the method is salient and unambiguous. The same methods may not, however, perform as well in real-life examples where different factors play a role. Therefore, one of the chapters in this dissertation (Chapter 4) is devoted to a comparative simulation study of several existing variance approximation methods in state space models. For this purpose, a real-world complex application is chosen: the Dutch Labour Force survey (DLFS) model that is used by Statistics Netherlands in the production of official statistical figures on the unemployed labour force. Apart from casting some light/shadow on these variance approximation methods, this study also suggests that simulating a model can be a good tool to check the model for overspecification. Such procedure suggests that it might be worth considering a more restricted version of the DLFS model. Biases in the Kalman filter-based variances in both variants of the DLFS model do not raise high concerns and become negligible as the time progresses, which makes the state space approach safe and attractive for use in the production of official figures. Another way to account for the uncertainty around (unknown) estimated hyperparameters is a full (hierarchical) Bayesian paradigm. Chapter 5 compares the state space approach with the multilevel Hierarchical Bayesian (HB) one in another official statistical application – the Dutch Transportation survey (DTS). It turns out that outcomes from the two approaches are quite comparable, with differences becoming visible mainly in small-scaled domains and applicable mainly to variance-, rather than point-estimates. Unlike in the DLFS, negative biases in standard error estimates produced by the Kalman filter are considerable in the case of the DTS reaching almost 8% at the provincial level and more than 10% at the national one. The more conservative standard error estimates from the multilevel HB model still offer a significant reduction in the design-based standard errors in the DTS: above 50% at the provincial level, and over 30% at the national level, averaged over the domains and time. In other words, in order to reduce the true variance within the design-based framework to the extent the time series techniques do, one would have to increase the sample size more than four-fold for the provincial level, and more than twice for the national level (conditional on the point-estimates). Chapter 5 also presents an approach for dealing with unreliable and missing design-based variance estimates that are used as input information in both multilevel and state space models. Many official statistical applications can be largely improved with the help of the time series techniques considered in this dissertation. At Statistics Netherlands, the US Census Bureau and the UK Office for National Statistics, certain attempts have been made to move further than the traditional sampling theory. As mentioned above, Statistics Netherlands is already producing official figures for the Labour Force survey based on the state space model. For the DTS presented in Chapter 5, certain attempts are being made to switch to the model-based times series approach in the production of official figures, as this survey is going to redesigned again in 2018. However, much more needs to be done in popularizing various modelling techniques among official statisticians. The applications presented in this dissertation have demonstrated how information accumulated over time and optionally domain space can be successfully used to increase the precision of survey estimates without resorting to (expensive) sample size increases, as well as in order to restore the comparability of figures before and after survey redesigns. The former must be of a particular importance at the time when most NSIs have to deal with reducing budgets. Examples of successful implementation of novel modelling methods and dissemination of these results is therefore crucial for advancement of official statistics. Hopefully, this dissertation provides practitioners with useful guidelines, and survey methodologists with sufficient evidence about the great potential of the time series techniques in repeatedly conducted surveys.